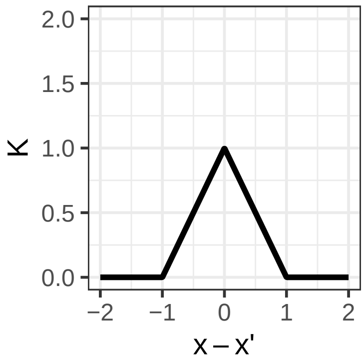
Solution 1: Gaussian Processes - Prediction

It may help to visualize the kernel function:



(a)

$$
\begin{aligned}
\mathbf{K} &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \text{(none of the training points are within 1 of each other)} \\
\boldsymbol{K}_*^T &= \begin{pmatrix} 0.6 \\ 0 \\ 0.3 \\ 0 \end{pmatrix}^T \\
m_{\text{post}} &= \begin{pmatrix} 0.6 \\ 0 \\ 0.3 \\ 0 \end{pmatrix}^T \begin{pmatrix} (1+\sigma^2)^{-1} & 0 & 0 & 0 \\ 0 & (1+\sigma^2)^{-1} & 0 & 0 \\ 0 & 0 & (1+\sigma^2)^{-1} & 0 \\ 0 & 0 & 0 & (1+\sigma^2)^{-1} \end{pmatrix} \boldsymbol{y} \\
&= \begin{pmatrix} \frac{0.6}{1+\sigma^2} & 0 & \frac{0.3}{1+\sigma^2} & 0 \end{pmatrix} \begin{pmatrix} 3 \\ 3.3 \\ 2.0 \\ 2.7 \end{pmatrix} \\
&= \frac{1.8}{1+\sigma^2} + \frac{0.6}{1+\sigma^2} \\
&= \frac{2.4}{1+\sigma^2}
\end{aligned}
$$

(b)

$$
\begin{aligned}
k_{\text{post}} &= 1 - \begin{pmatrix} \frac{0.6}{1+\sigma^2} & 0 & \frac{0.3}{1+\sigma^2} & 0 \end{pmatrix} \begin{pmatrix} 0.6 \\ 0 \\ 0.3 \\ 0 \end{pmatrix} \\
&= 1 - \left(\frac{0.36}{1+\sigma^2} \frac{0.09}{1+\sigma^2} \right) \\
&= 1 - \frac{0.45}{1+\sigma^2}
\end{aligned}
$$

(c)

$$m_{\text{post}} = \begin{pmatrix}1\\0\\0\\0\end{pmatrix}^T \begin{pmatrix}(1+\sigma^2)^{-1} & 0 & 0 & 0\\ 0 & (1+\sigma^2)^{-1} & 0 & 0\\ 0 & 0 & (1+\sigma^2)^{-1} & 0\\ 0 & 0 & 0 & (1+\sigma^2)^{-1}\end{pmatrix} \boldsymbol{y}$$

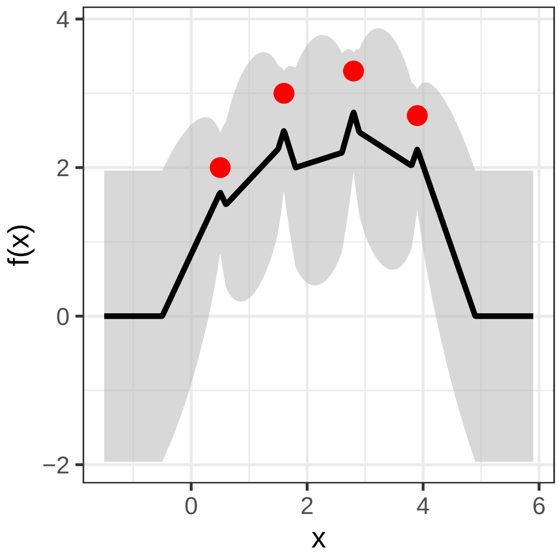
$$= \begin{pmatrix}(1+\sigma^2)^{-1} & 0 & 0 & 0\end{pmatrix} \boldsymbol{y}$$

$$= \frac{y^{(i)}}{1+\sigma^2}$$

$$k_{\text{post}} = 1 - \begin{pmatrix}1\\0\\0\\0\end{pmatrix}^T \begin{pmatrix}(1+\sigma^2)^{-1} & 0 & 0 & 0\\ 0 & (1+\sigma^2)^{-1} & 0 & 0\\ 0 & 0 & (1+\sigma^2)^{-1} & 0\\ 0 & 0 & 0 & (1+\sigma^2)^{-1}\end{pmatrix} \begin{pmatrix}1\\0\\0\\0\end{pmatrix}^T$$

$$= 1 - \frac{1}{1+\sigma^2} = \frac{\sigma^2}{1+\sigma^2}$$

(d) E.g., for $\sigma^2 = 0.2$:



(e) For $\sigma^2 = 0$, the "band" around the observed training points is smaller (the GP is noise-free):